ZooKeys 1270: 349–365 (2026)
DOI: 10.3897/zookeys.1270.170412

Research Article

Assessing fish diversity in small streams and ponds of the Peruvian Amazon using environmental DNA metabarcoding

Guillain Estivals[1,2], Ronald Delgado-Barboza[3], Morgan Ruiz-Tafur[2,4], Junior Chuctaya[2,4], Pierre Caminade[2,5], Carmen Garcia-Davila[1,2], Nicolas Hubert[2,6,7]

1 *Instituto de Investigaciones de la Amazonia Peruana (IIAP), AQUAREC, Laboratorio de Biología y Genética Molecular, Iquitos, Peru*
2 *Observatorio de la Biodiversidad de la Amazonia Peruana (OBAP), Instituto de Investigaciones de la Amazonia Peruana (IIAP), Avenida Quiñones Kilómetro 2.5, Iquitos, Loreto, Peru*
3 *Universidad Nacional San Antonio Abad del Cusco (UNSAAC), Peru*
4 *Instituto de Investigaciones de la Amazonia Peruana (IIAP), AQUAREC, Laboratorio de taxonomía de Peces, Iquitos, Peru*
5 *Centre National de la Recherche Scientifique, ISEM (Univ. Montpellier, CNRS, IRD), Montpellier, France*
6 *Institut de Recherche pour le Développement, ISEM (Univ. Montpellier, CNRS, IRD), Montpellier, France*
7 *Institut de Recherche pour le Développement (IRD), 44 boulevard de Dunkerque, Marseille, Bouches-du-Rhône, France*

Corresponding authors: Guillain Estivals (gestivals@iiap.gob.pe); Nicolas Hubert (nicolas.hubert@ird.fr); Junior Chuctaya (jchuctaya@iiap.gob.pe)

Abstract

The Amazon basin harbors exceptional fish diversity, with more than 3,500 species reported. However, this biodiversity is increasingly threatened by anthropogenic activities and climate change. The Peruvian Amazon alone is home to nearly 1,000 freshwater fish species – approximately one-third of the entire Amazon – yet significant gaps remain in our understanding of their distribution, ecology, and conservation status. Molecular approaches, particularly environmental DNA (eDNA) metabarcoding, have emerged as promising alternatives for rapid and accurate biodiversity assessment. In this study, a metabarcoding workflow targeting a fragment of the 12S gene to eDNA samples collected from small streams and ponds near Iquitos, Peru, was applied to evaluate the applicability of this approach for local fish community inventories. Water from 12 sites was filtered, and DNA was extracted, amplified, and sequenced using high-throughput Illumina technology. Bioinformatic analyses included MOTU clustering, haplotype identification, and taxonomic assignment using the Lowest Common Ancestor (LCA) algorithm. A total of 226 amplicon sequence variants (ASVs) were identified, 95 of which were assigned to fish taxa and clustered into 44 MOTUs across four orders: Characiformes, Gymnotiformes, Siluriformes, and Cichliformes. The study highlights the effectiveness of eDNA metabarcoding in detecting both common and elusive species, while also highlighting current limitations due to incomplete DNA barcode reference libraries. These findings underscore the need to expand barcode databases to fully harness the potential of eDNA for monitoring Amazonian fish biodiversity and support its use as a valuable tool for inventorying fish communities in sensitive and understudied environments.

Key words: 12S, ASV, biodiversity, JAMP, species delimitation, terra firme

OPEN ACCESS

Academic editor: Caleb McMahan
Received: 28 August 2025
Accepted: 16 January 2026
Published: 27 February 2026

ZooBank: https://zoobank.org/63805CE2-1505-4D32-9E6E-C051CAA849C1

Citation: Estivals G, Delgado-Barboza R, Ruiz-Tafur M, Chuctaya J, Caminade P, Garcia-Davila C, Hubert N (2026) Assessing fish diversity in small streams and ponds of the Peruvian Amazon using environmental DNA metabarcoding. ZooKeys 1270: 349–365. https://doi.org/10.3897/zookeys.1270.170412

Copyright: © Guillain Estivals et al.
This is an open access article distributed under terms of the Creative Commons Attribution License (Attribution 4.0 International – CC BY 4.0).

Introduction

With more than 3,500 fish species reported to date, the Amazon Basin is the most species-rich river system in the world (Jézéquel et al. 2020). However,

its biodiversity is among the most threatened (Malhi et al. 2008; Beuchle et al. 2021). Since the 1970s, anthropogenic activities, including deforestation, infrastructure development, mining, oil extraction, pollution, and overharvesting, have increased exponentially to meet the demands of a growing population and international markets (Beuchle et al. 2021; Albert et al. 2025). Moreover, the latest IPCC report highlights that temperatures will continue to rise in the Amazon (Lee et al. 2023), leading to major changes in water circulation across the continent (Espinoza et al. 2021). Climate change is currently driving major species reshuffling across the globe (Devictor et al. 2012; Pecl et al. 2017). However, these dynamics remain poorly documented in the Amazon due to significant gaps in biodiversity knowledge.

Aquatic ecosystems are keys components of the Amazon Basin (Castello et al. 2013; Albert et al. 2025), and fish are essential for local populations, as they represent a substantial source of animal protein (Adams et al. 2009). However, due to their high diversity and dramatic phenotypic changes during development, identifying fish species is a challenging task. In this context, taxonomic knowledge gaps related to species identity (i.e., the Linnean shortfall) and species distribution (i.e., the Wallacean shortfall) hinder accurate monitoring of Amazonian fish species. However, the development of molecular approaches, such as DNA barcoding for species delimitation and specimen identification, has greatly improved the accuracy and speed of ichthyological inventories in the Amazon, opening new perspectives for monitoring efforts (Pereira et al. 2013; Guimarães et al. 2018; Machado et al. 2018; de Freitas et al. 2024).

More than 1,000 fish species have been reported from the Peruvian freshwater ecosystems (Ortega et al. 2012) representing nearly one-third of all fish species found in the Amazon (Jézéquel et al. 2020). However, the range distribution, ecology, and conservation status of most of these species remain largely unknown. These major knowledge gaps are likely due to the complexity and vastness of the Amazon's aquatic environments, which make sampling individuals a challenging and time-consuming task. In this context, inventory methods adapted to these environments are required to accelerate research. Environmental DNA (eDNA) has opened new perspectives in this field, and metabarcoding – the simultaneous identification of numerous individuals through high-throughput sequencing – has the potential to become the fastest and most effective method for inventorying species across extensive geographical scales (Taberlet et al. 2018). One of its main strengths is its potential effectiveness to detect elusive species that are difficult to observe due to environmental conditions limiting visibility (such as currents, turbidity, and depth) or because of their ecological preferences (e.g., cryptobenthic species) (de Santana et al. 2021; Ruppert et al. 2019). To date, eDNA metabarcoding has increasingly been applied to assess fish diversity in the Amazon Basin, particularly using the 12S rRNA gene region (Cantera et al. 2019, 2022a, b, 2023; Cilleros et al. 2019; Bevilaqua et al. 2020; de Santana et al. 2021; Batista et al. 2022; Coutant et al. 2023; Condachou et al. 2024; Jackman et al. 2024; Martinelli Marín et al. 2024; Timana-Mendoza et al. 2025), the Cytochrome oxidase I gene (Mariac et al. 2022; Timana-Mendoza et al. 2025) and the 16S region (Bevilaqua et al. 2020), have also been explored. Most of these studies have shown that eDNA metabarcoding is a promising approach for inventorying the Amazonian ichthyofauna, despite reference libraries still being incomplete.

Here, we applied an eDNA metabarcoding workflow to samples collected in small streams and ponds near Iquitos, in the Peruvian Amazon. Using a fragment of the 12S gene region, we aimed to inventory fish communities and evaluate applicability of this method in this part of the basin. We further discuss the benefits, limitations, and implications of eDNA metabarcoding for improving biodiversity assessments in Amazonian freshwater ecosystem.

Materials and methods

Study area

The study was conducted in the Loreto Region of the Peruvian Amazon, around two towns: Jenaro Herrera on the right bank of the Ucayali River, and Nauta, on the left bank of the Marañón River (Fig. 1). Water samples were collected from ten sites in Jenaro Herrera area and two sites in the Nauta area. At each site, the highest volume possible was filtered, with an average volume of 22.5 L. However, standard deviation was high (SD ± 9.2 L) due to the occurrence of clogging. The sampling sites consisted of networks of smalls forest streams (~ 0.5–6 m wide) and ponds (~ 20–3,000 m^2) in terra firme forest. Water type across all sites ranged from clear to slightly blackish water, with no suspended sediment except at site 12, which showed a moderate sediment load.

eDNA sampling and extraction

Water samples from 12 sites were filtered using Waterra cartridge with a 0.45 µm pore size: ten sites in the region of Jenaro Herrera and two in the region of Nauta (Fig. 1). The filtration system consisted of tubing of 6 and 8 mm, and a water pump with a micro diaphragm (SEAFLO 12V 3,8 L/min, 40 psi) powered by a 12 V ion battery of a capacity of 7.5 mAh. To monitor potential contaminations related to pre-PCR steps, an additional cartridge was used in the field to filter 21 liters of bottled water. After filtering, the remaining water in each cartridge was expelled by introducing air into the system, and the cartridge was then filled (approx. 80 ml) with 1× Longmire buffer (Tris–HCl 0.1 M, EDTA 0.1 M, NaCl 0.01 M and N-lauroyl sarcosine 1% with pH 7.5–8). Each cartridge was labeled with a unique code and stored at 4 °C until DNA extraction.

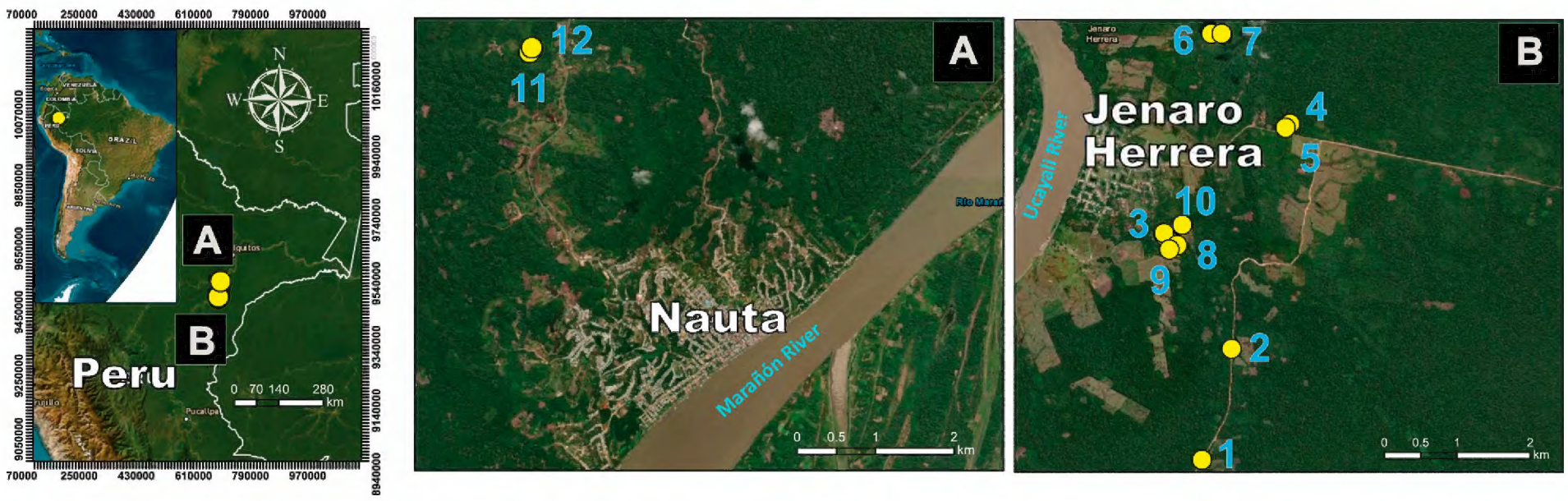

Figure 1. Map of the study area in the Loreto Department, Peru.

The eDNA contained within the cartridges was extracted using the following protocol: the cartridges were placed in a 56 °C incubator for 2 h. During the incubation, the cartridges were shaken at least twice. The contents of each cartridge were then poured and weighed into a biological sample collection bottle and divided equally into three 50-mL Oak Ridge High-Speed PPCO Centrifuge tubes (Nalgene™). The tubes were centrifuged at 15,000 × g for 15 min at 6 °C. After centrifugation, the supernatant from each tube was carefully pipetted off, leaving ~ 12 g of material in each tube. Twenty-seven milliliters of absolute ethanol and 1.2 mL of 3 M sodium acetate were added to the tubes, which were placed at -20 °C overnight. The tubes were subsequently centrifuged at 15,000 × g for 15 min at 6 °C, and the supernatants were carefully discarded by tilting. Next, 2 × 720 µL of ATL Buffer (Tissue lysis buffer, Qiagen) was added to each tube, and each was gently vortexed to dislodge the "clean" pellet from the bottom. The contents of each tube were carefully pipetted, avoiding the "dirty" fraction, and transferred into 2 mL tubes. To each tube, 20 µL of proteinase K (20 mg/µL) was added, and the samples were incubated for 2 hours at 56 °C in an Eppendorf ThermoMixer with shaking at 600 rpm. For each tube, 800 µL of the mixture was pipetted twice and distributed into two tubes of 2 mL, to which 250 µL of SB buffer was added. The tubes were vortexed and briefly centrifuged for ~ 5 s. The supernatants were then loaded onto columns corresponding to step 7 of the NucleoSpin Soil kit (Macherey-Nagel), and the extraction protocol was continued accordingly. In the final stage of the extraction, 50 µL of SE buffer pre-warmed to 56 °C was added on the membranes to elute the DNA. At the end of the extractions, 76 samples were obtained – six for each cartridge (6 × 13). To prevent cross-contamination, all extractions were conducted on a laboratory bench dedicated to eDNA analyses, sterilized with DNA away prior to each extraction. Pipettes were UV-sterilized before use, filtered tips were systematically used, and gloves were changed between each cartridge.

The quality and concentration of the DNA were measured using a nanodrop (thermos scientific), and the double-stranded DNA (dsDNA) were quantified with a Qubit 4 (Invitrogen).

DNA amplification, libraries construction and NGS sequencing

For each cartridge, three of the six aliquots with the highest dsDNA concentrations were selected for PCR amplification. Reactions were performed using AmpliTaq Gold 360 Master Mix (Applied Biosystem) and 12S Mifish primers (MiFish-F 5'_GTCGGTAAAACTCGTGCCAGC_3', MiFish-R 5'_CATAGTGGGGTATCTAATCCCAGTTTG_3'; Miya et al. 2015). An extension sequence was added to 5' end of both the forward and the reverse primers: TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG and GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG, respectively. Extensions were used for library indexing. Between the extension sequences and the primer sequence, one to three nucleotides were added to prevent signal saturation during Illumina Miseq sequencing. Each PCR reaction had a total volume of 20 µL and contained 10 µL of AmpliTaq Gold 360 Master Mix, 0.16 µL of bovine serum albumin (BSA), 5.84 µL of water, 2 µL combined forward and reverse primer mix (5 µM) and 2 µL of template eDNA. The program consisted of an initial Taq activation at 95 °C for 10 min, followed by 45 cycles of 95 °C for 30 s (denaturation), 60 °C for 30 s (annealing), and 72 °C for 60 s for extension, with a final extension step at 72 °C for 7 min.

Libraries were prepared using a two-step PCR protocol (Galan et al. 2018). In the first PCR step, a total of 168 samples were amplified, corresponding to each cartridge (12 sites + 1 negative control), three extraction aliquots, and four PCR replicates (13 cartridges × 12 replicates + 12 negative controls). In the second PCR step, each library was indexed using unique dual-index combinations to exclude chimeric amplicons more effectively during bioinformatics processing. Library was sequenced with 150 bp paired-end reads on an Illumina Miseq platform. The second PCR step and Illumina Miseq sequencing was performed by the Genseq platform (CNRS, Montpellier, France). Illumina Miseq sequencing run are available as SRA in NCBI GenBank (BioProject ID: PRJNA1391834).

Metabarcoding workflow

Bioinformatic analyses were performed on demultiplexed data using the JAMP pipeline developed by Elbrecht et al. (2018). This workflow uses the library "JAMP" in R, along with the tools VSEARCH (Rognes et al. 2016), USEARCH (Alloui et al. 2015) and CUTADAPT (Martin 2011). The main steps of the analysis included, reads merging, primers trimming, maximum expected error filtering, normalization to the same sequencing depth (to allow comparisons between sites), fragment length filtering, reads filtering to identify ASV (i.e., haplotype) including reads denoising, Molecular Operational Taxonomic Unit (MOTU) delimitation through ASV clustering based on a threshold approach (i.e., 97% similarity), and abundance-based filtering within and between libraries.

The sequence depth was normalized to 20000 reads per library, and fragment lengths were filtered to retain reads between 150 and 200 base pairs. No threshold filtering abundance was applied at the individual library level in order to be less conservative; instead, filtering was performed across libraries for each site (12 libraries per site). Haplotypes were retained if they were detected in at least six of the 12 libraries (≥50%) and discarded otherwise. Parameters used in the Denoise function were as follows: minsize = 20, minrelsize = 0, OTUmin = 0, minhaplosize = 0, withinOTU = 0, eachsampleOTUmin = 0, minHaploPresence = 6, minOTUPresence = 1, and other parameters were set to default values.

All haplotypes identified within each MOTU were compared against the NCBI-nr database using BLASTn (Chen et al. 2015), and the results were formatted using DIAMOND (Benson et al. 2013; Buchfink et al. 2015). BLASTn outputs were then used for taxonomic assignment with the Lowest Common Ancestor (LCA) algorithm implemented in Megan6 (Huson et al. 2016). In Megan, taxonomic identification is not based on the single best hit; instead, each read is assigned to the lowest taxonomic node that includes all valid hits above the defined thresholds (e.g., bit score, percent identity). Megan6 was run with default parameters and a minimum percent identity threshold of 98.0%; haplotype with blast identity scores below 98% were considered unassigned. ASVs collected here are made available in a fasta file in Suppl. material 3.

Spatial structure of fish communities

We examined the spatial structure of fish assemblages in the study area by performing a hierarchical cluster analysis based on the Raup-Crick dissimilarity index (Raup and Crick 1979). This index was preferred because it is a simple and

explicit way to measure dissimilarity in species composition between sites (0 = all species are shared, 0.5 = half of the species are shared, 1 = no species are shared) while adjusting for heterogeneous species richness between sites. The dissimilarity matrix among sites was calculated using the presence/absence of MOTUs using the raupcrick function from the R package vegan (Oksanen et al. 2022). The hierarchical clustering was then performed using the R package stats 3.1.2 (R Core Team 2025).

Results

Illumina Miseq sequencing generated a total of 5,864,056 reads across 168 libraries. From the 12 sampled sites, 226 ASVs were detected (Suppl. material 3), of which 95 could be assigned to the species level with the 98% threshold. Of these, all but one corresponded to fish species and were clustered into 44 MOTUs. Of these, 27 were identified at the species level, 12 at the genus level, one at the family level, two at the suborder level, and two at higher taxonomic ranks (Table 1, Fig. 2).

Table 1. Summary statistics of the similarity analysis for the 44 MOTUs assigned to a taxon.

Taxa	MOTU	Identity (%)	Nucleotide identity range	Nucleotide cover range	Total read number	Number of ASV
Specie level						
*Bujurquina mariae**	1	99	169	170	720	1
Crenicichla sp., *Crenicichla alta** (*Sexatilia alta**)	2	98–99	171–173	174	9008	1
Acestrorhynchus falcatus, *Acestrorhynchus* sp.	3	99	160–168	162–170	21733	3
*Leporinus lacustris**, *Leporinus* sp.	4	98–99	166–171	169–172	6231	6
*Leporinus piau**, *Leporinus* sp.	5	98–99	166–167	169	920	2
Schizodon fasciatus, *Schizodon* sp.	6	99–100	170–171	170–172	285	1
*Planaltina myersi**	7	98	103	105	638	1
Astyanax bimaculatus	8	98	98	100	5957	1
Tetragonopterus argenteus	9	99	172–173	174	1094	2
Crenuchus spilurus, *Crenuchus* sp.	10	99–100	170–172	172	7289	1
Erythrinus erythrinus, *Erythrinus* sp.	11	98–100	160–167	162–167	30846	7
Hoplerythrinus unitaeniatus, *Hoplerythrinus* sp.	12	99–100	163–166	165–166	43894	4
Hoplias malabaricus	13	98	163–164	166–167	10478	3
Carnegiella strigata	14	100	163–168	163–168	12382	2
*Prochilodus harttii**, *Prochilodus lineatus**, *Prochilodus costatus**, *Prochilodus nigricans*, *Prochilodus magdalenae**, *Prochilodus reticulatus**, *Prochilodus argenteus**	15	98–100	163–171	165–171	7049	1
Triportheus angulatus	16	100	170	170	642	1
*Electrophorus electricus**	17	100	167	167	1690	1
*Electrophorus voltai**, *Electrophorus* sp.	18	98–100	164–167	167	1138	1
Gymnotus carapo	19	98–100	167–172	170–172	42388	9
Apteronotus sp., *Apteronotus albifrons*	20	99	171	172	8397	1
Brachyhypopomus verdii, *Brachyhypopomus* sp.	21	98–99	164–167	165–168	9417	4
*Microsternarchus bilineatus**	22	98	170	173	504	1
Megalechis picta	23	98–99	174–175	177	654	1
*Ancistrus trinitatis**	24	99	169	171	2799	1
Hypostomus sp., *Hypostomus ancistroides**	25	98	168	171	683	1
Farlowella oxyrryncha, *Farlowella knerii*, *Farlowella schreitmuelleri**, *Farlowella paraguayensis**, *Farlowella hahni**, *Farlowella reticulata**, *Farlowella smithi*	26	98–100	170–173	173–174	1295	1
Rhamdia quelen	27	98–99	105–171	106–174	42940	2
Genus level						
***Aequidens*:** *Aequidens* sp.	29	98	162–167	165–170	24193	3

Taxa	MOTU	Identity (%)	Nucleotide identity range	Nucleotide cover range	Total read number	Number of ASV
***Hypselecara*:** *Hypselecara coryphaenoides** or *Hypselecara temporalis*	30	99	168	170	19004	3
***Crenicichla*:** *Crenicichla* sp., *Crenicichla alta** (*Sexatilia alta**), *Crenicichla frenata** (*Sexatilia frenata **)	31	98–99	168–171	171–174	2399	2
Moenkhausia sp.	32	99	166	168	7778	1
***Hoplias*:** *Hoplias malabaricus*, *Hoplias aimara**, *Hoplias* sp.	33	98–99	160–165	163–167	17168	2
Semaprochilodus sp.	34	98	169	172	398	1
***Brachyhypopomus*:** *Brachyhypopomus* sp.	35	99–100	166–168	169	3026	3
***Semaprochilodus*:** *Semaprochilodus* sp., *Semaprochilodus taeniurus*	36	99–100	169–172	170–172	14435	1
Eigenmannia sp.	37	98	173	176	6965	1
***Eigenmannia*:** *Eigenmannia* sp.	38	99	167–168	169	5181	2
Sternopygus sp.	39	98–99	165–172	168–174	42954	4
Peckoltia sp.	40	98	168	171	5578	2
Family level						
Serrasalmidae: *Myloplus rubripinnis**, *Myloplus* sp., *Myleus tiete*	41	99–100	172–173	173–175	59667	1
Suborder						
Characoidei: *Metynnis maculatus*, *Brycon melanopterus*	42	99–100	166–168	168	6253	3
Siluroidei: *Pimelodella* sp., *Pimelodus mysteriosus**, *Pimelodus blochii*, *Pimelodus* cf. *blochii*, *Pimelodus pohli**, *Pimelodus yuma**, *Pimelodus argenteus**, *Pimelodus coprophagus**, *Pimelodus* sp.	43	98–99	168–172	171–175	710	1
Subdivision						
Percomorphaceae: *Australoheros facetus**, *Aequidens metae**, *Pomadasys ramosus**, *Centropomus undecimalis**, *Bairdiella ronchus**, *Centropomus parallelus**, *Pachyurus adspersus**, *Micropogonias furnieri**, *Cichlasoma* sp.	44	98–100	160–169	163–170	97410	3
Supercohort						
Clupeocephala: *Serrasalmus* sp., *Pygocentrus nattereri*, *Pterophyllum scalare*, *Pygocentrus piraya**, *Serrasalmus rhombeus*, *Serrasalmus brandtii**, *Serrasalmus serrulatus*, *Serrasalmus eigenmanni*, *Serrasalmus elongatus*, *Pristobrycon striolatus*	45	98–100	160–170	162–171	8201	1
Not assigned						131

Note: Taxa identified with an asterisk (*) are not known to be distributed in Peru.

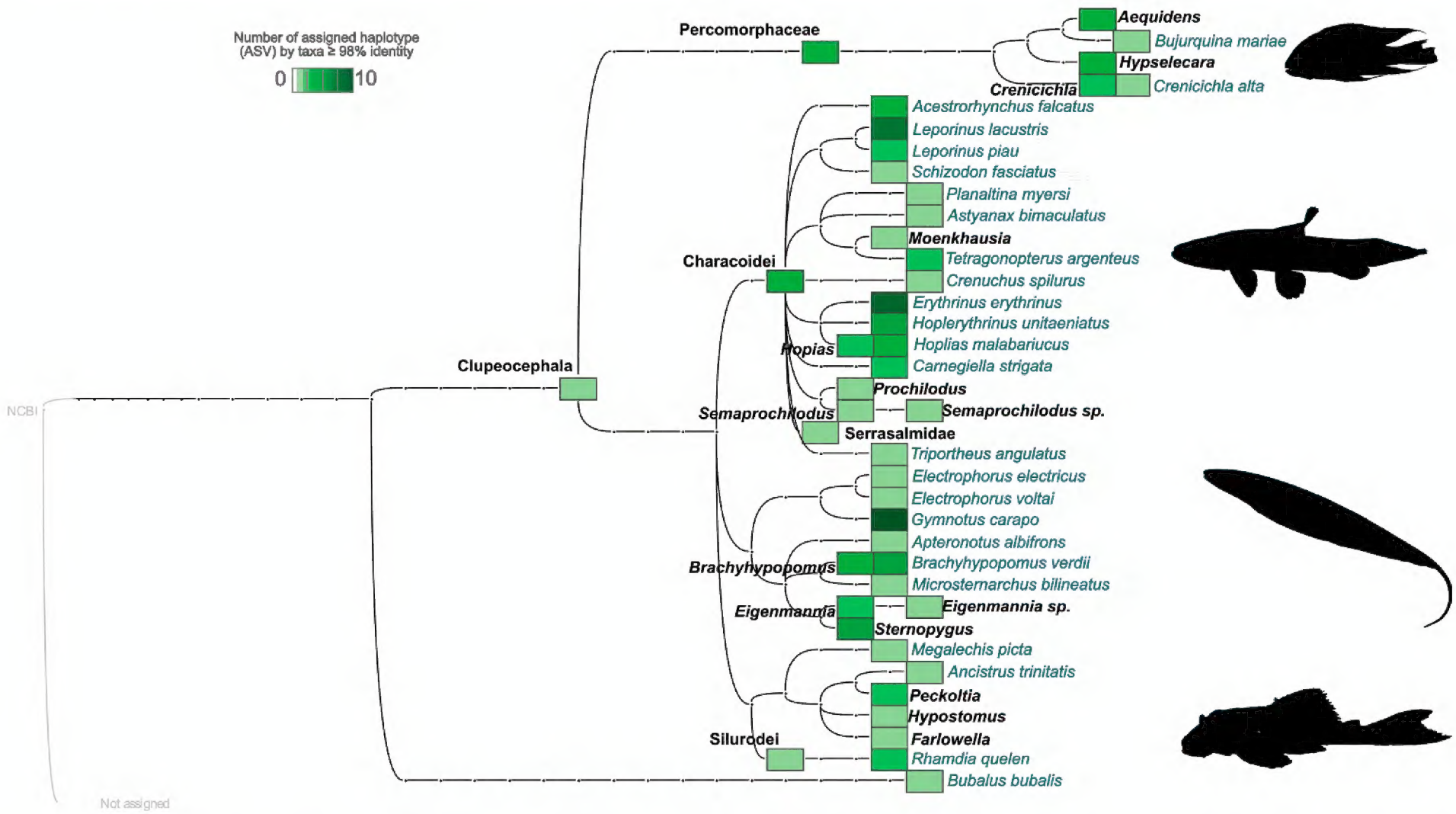

Figure 2. Taxonomic assignments using the LCA algorithm with default parameters and setting the minimum percent identity to 98.0% with MEGAN6.

The 44 MOTUs were assigned to four orders: Characiformes (45%), Gymnotiformes (23%), Siluriformes (16%), and Cichliformes (14%), with an additional 2% corresponding to an unresolved basal node (Fig. 2). The number of MOTUs detected per site ranged from one (site 7) to 19 (site 12). The most frequently detected species across sites was *Gymnotus carapo*, found in five of 12 sites. In terms of reads counts, *Hoplerythrinus unitaeniatus* (43,894 reads), *Rhamdia quelen* (42,940 reads), and *Gymnotus carapo* (42,388 reads) were the most abundant taxa (Table 1).

The species with the highest number of ASVs were *Gymnotus carapo* (9 ASVs), *Erythrinus erythrinus* (7 ASVs), and *Leporinus piau* (6 ASVs) (Suppl. material 1). For *Gymnotus carapo*, four ASVs were found exclusively at site 6, one at site 1, and one ASV was shared between sites 9 and 11. For *Erythrinus erythrinus*, six ASVs were found at site 11, three at site 1, one ASV was shared among sites 1, 2, 10, and 11, and another ASV was shared between sites 1 and 11. All six ASVs identified for *Leporinus lacustris* were found at site 12.

Site 12, in addition to being the richest in terms of species among all the studied sites, exhibited the most dissimilar fish community (Fig. 3). The other sites

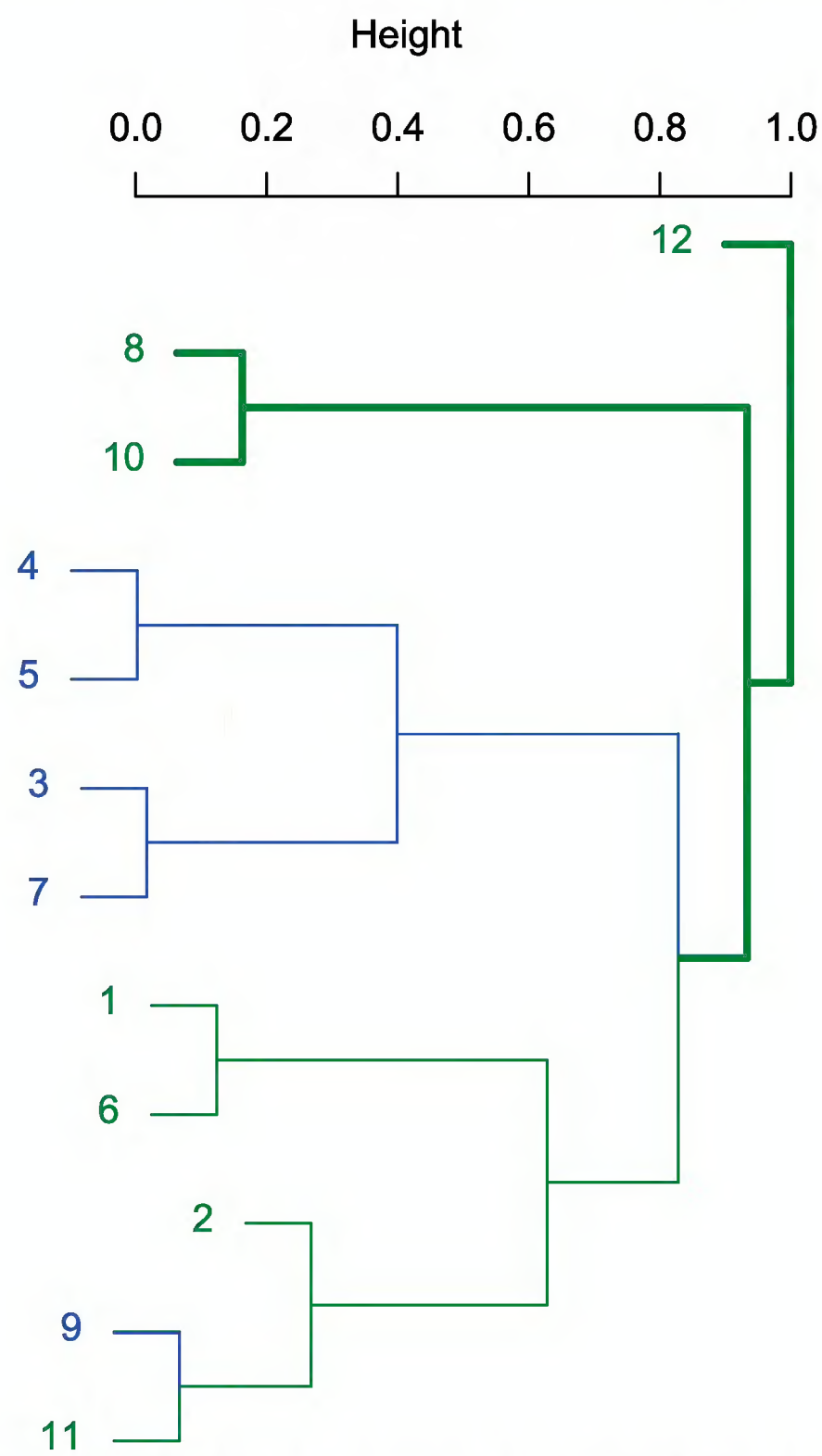

Figure 3. Hierarchical cluster of the 12 sampled sites using the Raup-Crick dissimilarity index and presence/absence of MOTUs, green lines represent rivers (largest lines) and streams (smaller lines), blue lines represent ponds.

are clustered into three main groups: one cluster comprising sites 8 and 10, another one with sites 3, 4, 5, and 7, and a final cluster with sites 1, 2, 6, 9, and 11 (Fig. 3). Sites 8, 10, and 12 correspond to the largest tributaries of the 12 sites, measuring approximately 3 to 6 meters wide, with site 10 being a tributary of site 8. Sites 3, 4, 5, and 7 correspond to ponds; sites 4 and 5 were sampled from the same pond. Except for site 9, which is the smallest pond sampled, sites 1, 2, 6 and 11 correspond to the smallest streams, each less than 1 m wide.

Discussion

To our knowledge, this is the first study to apply environmental DNA metabarcoding to inventory fish communities specifically in small terra firme streams (<1 m wide) and ponds (20–3,000 m^2) of the Peruvian Amazon. Across all sites studied, we observed a dominance, in term of number of species, belonging to the orders Characiformes, Gymnotiformes, and Siluriformes. These observations differ slightly from the overall Amazonian ichthyofauna, which is mainly represented by the orders Characiformes, Siluriformes and Cichliformes (Dagosta and De Pinna 2019). Other authors have reported different results based on samplings methods using both individual collections and eDNA (Zuanon et al. 2015; Jackman et al. 2021; Batista et al. 2022). These differences can be explained by the uneven distribution of the orders that make up the Amazonian ichthyofauna throughout the watershed. Variation in distribution is primarily influenced by the geological and hydrological history of the Amazon basin (Cassemiro et al. 2023; Hubert and Renno 2006), as well as the habitat heterogeneity and the ecological adaptations or preferences of certain groups to specific environments (Oberdorff et al. 2019). However, it is interesting to note the omnipresence and dominance of Characiformes in Amazonian environments, even in small streams. The sites sampled were forest streams on terra firme, characterized by low conductivity and a pH below 7. It has been shown that several genera of Gymnotiformes are associated with such environments, including *Gymnotus*, *Brachyhypopomus*, *Hypopygus*, *Microsternarchus*, *Gymnorhamphichthys*, *Sternopygus* and *Eigenmannia* (Albert and Crampton 2005). In our study, we detected all these genera except *Hypopygus* and *Gymnorhamphichthys*, as well as *Electrophorus* and *Apteronotus*.

In general, the genera and species identified in this study are known to occur in Peru and are expected in this type of environment. However, several species detected in our dataset – such as *Leporinus lacustris*, *Leporinus piau*, *Planaltina myersi*, *Electrophorus electricus*, *Electrophorus voltai* and *Ancistrus trinitatis* – have not previously been recorded in Peru. Some of these records may reflect recent taxonomic revisions (e.g., *Electrophorus*), but they are more attributable to misidentifications in the NCBI database and potential errors in the phylogenetic assignments performed by MEGAN during Lowest Common Ancestor (LCA) classification.

Additional factors contributing to uncertain assignments include the absence of 12S barcodes for many local species and the low genetic divergence among congeneric taxa. This is well exemplified by the spurious detection of *Ancistrus trinitatis*, a species originally described from the island of Trinidad and belonging to the taxonomically complex genus *Ancistrus*, which includes numerous closely related and poorly defined species. Similarly, *Leporinus piau*

and *L. lacustris* were detected in our dataset, although their known distributions are restricted to northeastern Brazilian coastal rivers and the Paraná River Basin, respectively. These improbable records most likely reflect incomplete or misannotated reference libraries, a common limitation in Amazonian eDNA studies. Taken together, these observations emphasize the need for expanding and curating regional molecular reference databases to improve species-level resolution and underscore the importance of interpreting eDNA-based taxonomic assignments with caution (Blackman et al. 2023; Gold et al. 2021; Keck et al. 2023). For instance, of the 226 ASVs identified, 131 could not be assigned to any taxon with an identity ≥ 98% (Suppl. material 2), but were assigned to a sequence with ≥ 98%. This was due to the absence of 12S reference barcodes for these particular taxa.

Besides, the Linnean shortfall – i.e., the gap between described species and actual biodiversity – further limit the effectiveness of eDNA metabarcoding, which remains constrained to a relatively small set of well-known species. According to Marques et al. (2021) and the GAPeDNA v. 1.1.2 platform (https://shiny.cefe.cnrs.fr/GAPeDNA/), only 11% (254 of 2273) of Amazonian fish species have a 12S reference barcode (MiFish), and 40–60% of species classified as vulnerable by the IUCN lack any reference sequence. This lack of 12S reference information is likely one of the main reasons for the relatively low richness of Siluriformes recorded in our study. Many siluriform species occurring in the Amazon probably lack 12S barcode data, which limits their detection and assignment in metabarcoding analyses.

During our study, conducted as part of a project assessing the reliability of eDNA for characterizing intraspecific genetic diversity of *Apistogramma agassizii*, individuals of *A. agassizii* were collected from 11 of the 12 sites. However, assignments with ≥98% identity did not confirm the presence of *A. agassizii* in the study area. Among the 131 unassigned ASVs, some were attributed to *A. agassizii*, but the identity percentages did not exceed 94%. Sanger sequencing of the 12S fragment of *A. agassizii* individuals collected in the study area (*A. agassizii* Sp1, cf.: Estivals et al. 2023) confirms the presence of the species, matching with 100% identity to one of the previously unassigned ASVs. This example perfectly illustrates the impact of taxonomic gaps in the reference libraries and highlights the need to generate reference barcodes to fill those gaps and ensure the full effectiveness of eDNA metabarcoding approaches.

Finally, our study supports previous findings suggesting that eDNA effectively captures the structure of fish assemblages in the Amazon (Cantera et al. 2022a; Mariac et al. 2022), offering some new perspectives for the characterization of the Amazonian biotas. The sites corresponding to the largest tributaries (8, 10, and 12) host the highest number of species – an observation previously reported by Stegmann et al. (2019). However, the opposite trend is observed at the scale of the Amazon basin due to the complex geological and hydrographical history of the basin (Carvajal-Quintero et al. 2019; Oberdorff et al. 2019). At the more local scale of our study, larger rivers likely support higher species richness because they provide more space, a greater variety of habitats, and increased food availability. Interestingly, fish communities in our study also grouped according to environmental type – lotic versus lentic systems – consistent with patterns observed at broader scales (Mariac et al. 2022).

Conclusions

Our study demonstrates that the fragment of the 12S gene amplified with the MiFish primers is highly specific and effective for inventorying the Amazonian ichthyofauna. The detection of numerous ASVs (haplotypes) within certain species suggests that the 12S MiFish fragment may also be valuable for intraspecific genetic studies. Overall, eDNA metabarcoding presents an interesting alternative to traditional individual sampling, especially in sensitive environments such as small rivers in the Amazon. However, as highlighted by most studies, a significant gap remains in the barcode reference libraries. The main challenge in the coming years will not be the technical ability to conduct studies eDNA metabarcoding studies in the Amazon, but rather to bridge the barcode gaps across Amazonian countries, which currently limit the full effectiveness of eDNA approaches.

Additional information

Conflict of interest

The authors have declared that no competing interests exist.

Ethical statement

No ethical statement was reported.

Use of AI

No use of AI was reported.

Funding

This work was supported by the CONCYTEC through a scholarship to Guillain Estivals (contract [N° PE501083194-2023-PROCIENCIA]) from the PROCIENCIA fundings scheme " Proyectos de Investigación Básica 2023-01 E041-2023-01 " and funding from the Observatorio de la Biodiversidad de la Amazonia Peruana (OBAP). Data presented in this publication were acquired in the framework of OBAP (Observatorio de la Biodiversidad de la Amazonia Peruana), jointly developed by IIAP and IRD.

Author contributions

Conceptualization: GE, PC, CGD and NH. Data curation: GE and RJDB. Formal analysis: GE, RJDB and NH. Funding acquisition: GE. Investigation: GE, RJDB, MRT, JC, PC, CGD and NH. Methodology: GE, PC and NH. Project administration: GE. Software: GE. Supervision: GE and NH. Validation: GE, RJDB, MRT, JC, PC, CGD and NH. Visualization: GE. Writing – original draft: GE and NH. Writing – review and editing: GE, RJDB, MRT, JC, PC and CGD.

Author ORCIDs

Guillain Estivals https://orcid.org/0000-0001-5568-3548
Morgan Ruiz-Tafur https://orcid.org/0000-0002-8412-7564
Junior Chuctaya https://orcid.org/0000-0002-8876-4675
Carmen Garcia-Davila https://orcid.org/0000-0003-4125-5563
Nicolas Hubert https://orcid.org/0000-0001-9248-3377

Data availability

All of the data that support the findings of this study are available in the main text or Supplementary Information.

References

Adams C, Murrieta R, Neves W, Harris M (2009) Amazon Peasant Societies in a Changing Environment: Political Ecology, Invisibility and Modernity in the Rainforest. Springer, Netherlands, 358 pp. https://doi.org/10.1007/978-1-4020-9283-1

Albert JS, Crampton WGR (2005) Diversity and Phylogeny of Neotropical Electric Fishes (Gymnotiformes). Electroreception. Springer, 360–409. https://doi.org/10.1007/0-387-28275-0_13

Albert JS, Carnaval AC, Flantua SGA, Lohmann LG, Ribas CC, Riff D, Carrillo JD, Fan Y, Figueiredo JJP, Guayasamin JM, Hoorn C, de Melo GH, Nascimento N, Quesada CA, Ulloa Ulloa C, Val P, Arieira J, Encalada AC, Nobre CA (2025) Human impacts outpace natural processes in the Amazon. Science 379(6630): eabo5003. https://doi.org/10.1126/science.abo5003

Alloui T, Bousseboughl, Chaoui A, Nouar AZ, Chettah MC (2015) Usearch: A Meta Search Engine based on a new result merging strategy. 2015 7th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (IC3K), 531–536. https://doi.org/10.5220/0005642905310536

Batista LM, de Sá-Leitão CS, de Souza ÉMS, dos Anjos-Santos CH, de Almeida-Val VMF (2022) Addressing amazonian fish diversity using environmental DNA (eDNA): A first glance. European Journal of Aquatic Sciences 1(1): 9–17. https://doi.org/10.24018/ejaqua.2022.1.1.4

Benson DA, Cavanaugh M, Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW (2013) GenBank. Nucleic Acids Research 41(D1): D36–D42. https://doi.org/10.1093/nar/gks1195

Beuchle R, Achard F, Bourgoin C, Vancutsem C, Eva H, Follador M (2021) Deforestation and Forest degradation in the Amazon. European Union, Luxembourg. https://doi.org/10.2760/61682 [online]

Bevilaqua DR, de Melo SA, de Carvalho Freitas CE, da Silva ACV, da Silva Batista J (2020) First environmental DNA (eDNA) record of central Amazon in a floodplain lake: extraction method selection and validation. Brazilian Journal of Development 6: 87606–87621. https://doi.org/10.34117/bjdv6n11-254

Blackman RC, Walser J, Rüber L, Brantschen J, Villalba S, Brodersen J, Seehausen O, Altermatt F (2023) General principles for assignments of communities from eDNA: Open versus closed taxonomic databases. Environmental DNA 5(2): 326–342. https://doi.org/10.1002/edn3.382

Buchfink B, Xie C, Huson DH (2015) Fast and sensitive protein alignment using DIAMOND. Nature Methods 12(1): 59–60. https://doi.org/10.1038/nmeth.3176

Cantera I, Cilleros K, Valentini A, Cerdan A, Dejean T, Iribar A, Taberlet P, Vigouroux R, Brosse S (2019) Optimizing environmental DNA sampling effort for fish inventories in tropical streams and rivers. Scientific Reports 9(1): e3085. https://doi.org/10.1038/s41598-019-39399-5

Cantera I, Decotte J, Dejean T, Murienne J, Vigouroux R, Valentini A, Brosse S (2022a) Characterizing the spatial signal of environmental DNA in river systems using a community ecology approach. Molecular Ecology Resources 22(4): 1274–1283. https://doi.org/10.1111/1755-0998.13544

Cantera I, Coutant O, Jézéquel C, Decotte J-B, Dejean T, Iribar A, Vigouroux R, Valentini A, Murienne J, Brosse S (2022b) Low level of anthropization linked to harsh vertebrate biodiversity declines in Amazonia. Nature Communications 13(1): e3290. https://doi.org/10.1038/s41467-022-30842-2

Cantera I, Jézéquel C, Dejean T, Murienne J, Vigouroux R, Valentini A, Brosse S (2023) Functional responses to deforestation in fish communities inhabiting neotropical streams and rivers. Ecological Processes 12(1): 1–52. https://doi.org/10.1186/s13717-023-00463-8

Carvajal-Quintero J, Villalobos F, Oberdorff T, Grenouillet G, Brosse S, Hugueny B, Jézéquel C, Tedesco PA (2019) Drainage network position and historical connectivity explain global patterns in freshwater fishes' range size. Proceedings of the National Academy of Sciences of the United States of America 116(27): 13434–13439. https://doi.org/10.1073/pnas.1902484116

Cassemiro FAS, Albert JS, Antonelli A, Menegotto A, Wüest RO, Cerezer F, Coelho MTP, Reis RE, Tan M, Tagliacollo V, Bailly D, da Silva VFB, Frota A, da Graça WJ, Ré R, Ramos T, Oliveira AG, Dias MS, Colwell RK, Rangel TF, Graham CH (2023) Landscape dynamics and diversification of the megadiverse South American freshwater fish fauna. Proceedings of the National Academy of Sciences of the United States of America 120(2): e2211974120. https://doi.org/10.1073/pnas.2211974120

Castello L, McGrath DG, Hess LL, Coe MT, Lefebvre PA, Petry P, Macedo MN, Renó VF, Arantes CC (2013) The vulnerability of Amazon freshwater ecosystems. Conservation Letters 6: 217–229. https://doi.org/10.1111/conl.12008

Chen Y, Ye W, Zhang Y, Xu Y (2015) High speed BLASTN: An accelerated MegaBLAST search tool. Nucleic Acids Research 43(16): 7762–7768. https://doi.org/10.1093/nar/gkv784

Cilleros K, Valentini A, Allard L, Dejean T, Etienne R, Grenouillet G, Iribar A, Taberlet P, Vigouroux R, Brosse S (2019) Unlocking biodiversity and conservation studies in high-diversity environments using environmental DNA (eDNA): A test with Guianese freshwater fishes. Molecular Ecology Resources 19(1): 27–46. https://doi.org/10.1111/1755-0998.12900

Condachou C, Coutant O, Brosse S, Murienne J (2024) Evaluating the eDNA spatial signal in natural populations and its application for the monitoring of the threatened *Harttiella* (Siluriformes, Loricariidae). Limnology and Oceanography: Methods 22(9): 660–669. https://doi.org/10.1002/lom3.10634

Coutant O, Jézéquel C, Mokany K, Cantera I, Covain R, Valentini A, Dejean T, Brosse S, Murienne J (2023) Environmental DNA reveals a mismatch between diversity facets of Amazonian fishes in response to contrasting geographical, environmental and anthropogenic effects. Global Change Biology 29(7): 1741–1758. https://doi.org/10.1111/gcb.16533

Dagosta FCP, De Pinna M (2019) The fishes of the Amazon: Distribution and biogeographical patterns, with a comprehensive list of species. Bulletin of the American Museum of Natural History 2019(431): 1–163. https://doi.org/10.1206/0003-0090.431.1.1

de Freitas EA, Dos Santos DB, Moraes Ferreira CS, Silva-Oliveira C, Evangelista-Gomes GF, Veneza IB (2024) Integrative use of DNA barcode and morphology reveal high level of diversity in the ornamental fish on the lower Amazon basin. PLoS ONE 19(12): e0316455. https://doi.org/10.1371/journal.pone.0316455

de Santana CD, Parenti LR, Dillman CB, Coddington JA, Bastos DA, Baldwin CC, Zuanon J, Torrente-Vilara G, Covain R, Menezes NA, Datovo A, Sado T, Miya M (2021) The critical role of natural history museums in advancing eDNA for biodiversity studies:

A case study with Amazonian fishes. Scientific Reports 11(1): e18159. https://doi.org/10.1038/s41598-021-97128-3

Devictor V, Van Swaay C, Brereton T, Brotons L, Chamberlain D, Heliölä J, Herrando S, Julliard R, Kuussaari M, Lindström Å (2012) Differences in the climatic debts of birds and butterflies at a continental scale. Nature Climate Change 2: 121–124. https://doi.org/10.1038/nclimate1347

Elbrecht V, Vamos EE, Steinke D, Leese F (2018) Estimating intraspecific genetic diversity from community DNA metabarcoding data. PeerJ 6: e4644. https://doi.org/10.7717/peerj.4644

Espinoza J-C, Arias PA, Moron V, Junquas C, Segura H, Sierra-Perez JP, Wongchuig S, Condom T (2021) Impacts on Precipitation and Fire Season. Journal of Climate 34: 9025–9042. https://www.jstor.org/stable/27202166

Estivals G, Duponchelle F, García-Dávila C, Römer U, Mariac C, Renno J-F (2023) Exceptional Genetic Differentiation at a Micro-geographic Scale in Apistogramma agassizii (Steindachner, 1875) from the Peruvian Amazon: Sympatric Speciation? Evolutionary Biology 50(1): 1–17. https://doi.org/10.1007/s11692-022-09587-9

Galan M, Pons J, Tournayre O, Pierre E, Leuchtmann M, Pontier D, Charbonnel N (2018) Metabarcoding for the parallel identification of several hundred predators and their prey: Application to bat species diet analysis. Molecular Ecology Resources 18(3): 474–489. https://doi.org/10.1111/1755-0998.12749

Gold Z, Curd EE, Goodwin KD, Choi ES, Frable BW, Thompson AR, Walker Jr HJ, Burton RS, Kacev D, Martz LD, Barber PH (2021) Improving metabarcoding taxonomic assignment: A case study of fishes in a large marine ecosystem. Molecular Ecology Resources 21(7): 2546–2564. https://doi.org/10.1111/1755-0998.13450

Guimarães KLA, de Sousa MPA, Ribeiro FRV, Porto JIR, Rodrigues LRR (2018) DNA barcoding of fish fauna from low order streams of Tapajós River basin. PLoS ONE 13(12): e0209430. https://doi.org/10.1371/journal.pone.0209430

Hubert N, Renno J (2006) Historical biogeography of South American freshwater fishes. Journal of Biogeography 33(8): 1414–1436. https://doi.org/10.1111/j.1365-2699.2006.01518.x

Huson DH, Beier S, Flade I, Górska A, El-Hadidi M, Mitra S, Ruscheweyh H-J, Tappu R (2016) MEGAN Community Edition – Interactive Exploration and Analysis of Large-Scale Microbiome Sequencing Data. PLoS Computational Biology 12(6): e1004957. https://doi.org/10.1371/journal.pcbi.1004957

Jackman JM, Benvenuto C, Coscia I, Oliveira Carvalho C, Ready JS, Boubli JP, Magnusson WE, McDevitt AD, Guimarães Sales N (2021) eDNA in a bottleneck: Obstacles to fish metabarcoding studies in megadiverse freshwater systems. Environmental DNA 3(4): 837–849. https://doi.org/10.1002/edn3.191

Jackman JM, Sales NG, Benvenuto C, Drewitt A, Wolfenden A, Robins P, Coscia I, McDevitt A (2024) Performance of eDNA capture methods for monitoring fish biodiversity in a hyper-tidal estuary. bioRxiv: 2012–2024. https://doi.org/10.1101/2024.12.20.629404

Jézéquel C, Tedesco PA, Bigorne R, Maldonado-Ocampo JA, Ortega H, Hidalgo M, Martens K, Torrente-Vilara G, Zuanon J, Acosta A, Agudelo E, Barrera Maure S, Bastos DA, Bogotá Gregory J, Cabeceira FG, Canto ALC, Carvajal-Vallejos FM, Carvalho LN, Cella-Ribeiro A, Covain R, Donascimiento C, Dória CRC, Duarte C, Ferreira EJG, Galuch AV, Giarrizzo T, Leitão RP, Lundberg JG, Maldonado M, Mojica JI, Montag LFA, Ohara WM, Pires THS, Pouilly M, Prada-Pedreros S, de Queiroz LJ, Rapp Py-Daniel L, Ribeiro FRV, Ríos Herrera R, Sarmiento J, Sousa LM, Stegmann LF, Valdiviezo-Rivera J, Villa

F, Yunoki T, Oberdorff T (2020) A database of freshwater fish species of the Amazon Basin. Scientific Data 7(1): e96. https://doi.org/10.1038/s41597-020-0436-4

Keck F, Couton M, Altermatt F (2023) Navigating the seven challenges of taxonomic reference databases in metabarcoding analyses. Molecular Ecology Resources 23(4): 742–755. https://doi.org/10.1111/1755-0998.13746

Lee H, Calvin K, Dasgupta D, Krinner G, Mukherji A, Thorne P, Trisos C, Romero J, Aldunce P, Barret K (2023) IPCC, 2023: Climate Change 2023: Synthesis Report, Summary for Policymakers. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, H. Lee and J. Romero (eds.)]. IPCC, Geneva, Switzerland. https://www.ipcc.ch/report/ar6/syr/downloads/report/IPCC_AR6_SYR_SPM.pdf

Machado VN, Collins RA, Ota RP, Andrade MC, Farias IP, Hrbek T (2018) One thousand DNA barcodes of piranhas and pacus reveal geographic structure and unrecognised diversity in the Amazon. Scientific Reports 8(1): e8387. https://doi.org/10.1038/s41598-018-26550-x

Malhi Y, Roberts JT, Betts RA, Killeen TJ, Li W, Nobre CA (2008) Climate change, deforestation, and the fate of the Amazon. Science 319: 169–172. https://doi.org/10.1126/science.1146961

Mariac C, Duponchelle F, Miranda G, Ramallo C, Wallace R, Tarifa G, Garcia-Davila C, Ortega H, Pinto J, Renno J-F (2022) Unveiling biogeographical patterns of the ichthyofauna in the Tuichi basin, a biodiversity hotspot in the Bolivian Amazon, using environmental DNA. PLoS ONE 17(1): e0262357. https://doi.org/10.1371/journal.pone.0262357

Marques V, Milhau T, Albouy C, Dejean T, Manel S, Mouillot D, Juhel J (2021) GAPeDNA: Assessing and mapping global species gaps in genetic databases for eDNA metabarcoding. Diversity & Distributions 27(10): 1880–1892. https://doi.org/10.1111/ddi.13142

Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal 17: 10–12. https://doi.org/10.14806/ej.17.1.200

Martinelli Marín D, Lasso CA, Caballero Gaitan SJ (2024) eDNA metabarcoding: An effective tool for vertebrate diversity studies in the Colombian Amazon and Orinoco basins. Frontiers in Ecology and Evolution 12: e1409296. https://doi.org/10.3389/fevo.2024.1409296

Miya M, Sato Y, Fukunaga T, Sado T, Poulsen JY, Sato K, Minamoto T, Yamamoto S, Yamanaka H, Araki H, Kondoh M, Iwasaki W (2015) MiFish, a set of universal PCR primers for metabarcoding environmental DNA from fishes: Detection of more than 230 subtropical marine species. Royal Society Open Science 2(7): e150088. https://doi.org/10.1098/rsos.150088

Oberdorff T, Sversut Dias M, Jézéquel C, Albert J, Arantes C, Bigorne R, Carvajal-Valleros F, De Wever A, Frederico R, Hidalgo M, Hugueny B, Leprieur F, Maldonado M, Maldonado Ocampo J, Martens K, Ortega H, Sarmiento J, Tedesco P, Torrente-Vilara G, Zuanon J (2019) Unexpected fish diversity gradients in the Amazon basin. Science Advances 5(9): eaav8681. https://doi.org/10.1126/sciadv.aav8681

Oksanen J, Simpson G, Blanchet FG, Kindt R, Legendre P, Minchin P, O'Hara R, Solymos P, Szöcs E, Wagner H, Barbour M, Bedward M, Bolker B, Borcard D, Carvalho G, Chirico M, De Cáceres M, Durand S, Weedon J (2022) Vegan Community Ecology Package Version 2.6-2 April 2022. http://CRAN.R-project.org/package=vegan

Ortega H, Hidalgo M, Trevejo G, Correa E, Cortijo A, Meza-Vargas V, Espino J (2012) Lista anotada de los peces de aguas continentales del Perú: Estado actual del con-

ocimiento, distribución, usos y aspectos de conservación. Ministerio del Ambiente, Lima, 56 pp. https://museohn.unmsm.edu.pe/docs/pub_ictio/Ortega_et_al.2012Lista_Peces_Aguas_Cont.Peru.pdf

Pecl G, Araújo M, Bell J, Blanchard J, Bonebrake T, Chen I-C, Clark T, Colwell R, Danielsen F, Evengård B, Falconi L, Ferrier S, Frusher S, Garcia R, Griffis R, Hobday A, Janion C, Jarzyna M, Jennings S, Williams S (2017) Biodiversity redistribution under climate change: Impacts on ecosystems and human well-being. Science 355. https://doi.org/10.1126/science.aai9214

Pereira LHG, Hanner R, Foresti F, Oliveira C (2013) Can DNA barcoding accurately discriminate megadiverse Neotropical freshwater fish fauna? BMC Genetics 14(1): 1–20. https://doi.org/10.1186/1471-2156-14-20

R Core Team (2025) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna. https://www.R-project.org/

Raup DM, Crick RE (1979) Measurement of Faunal Similarity in Paleontology. Journal of Paleontology 53: 1213–1227. http://www.jstor.org/stable/1304099

Rognes T, Flouri T, Nichols B, Quince C, Mahé F (2016) VSEARCH: A versatile open source tool for metagenomics. PeerJ 4: e2584. https://doi.org/10.7717/peerj.2584

Ruppert KM, Kline RJ, Rahman MS (2019) Past, present, and future perspectives of environmental DNA (eDNA) metabarcoding: A systematic review in methods, monitoring, and applications of global eDNA. Global Ecology and Conservation 17: e00547. https://doi.org/10.1016/j.gecco.2019.e00547

Stegmann LF, Leitao RP, Zuanon J, Magnusson WE (2019) Distance to large rivers affects fish diversity patterns in highly dynamic streams of Central Amazonia. PLoS ONE 14(10): e0223880. https://doi.org/10.1371/journal.pone.0223880

Taberlet P, Bonin A, Zinger L, Coissac E (2018) Environmental DNA: For Biodiversity Research and Monitoring. Oxford University Press, 253 pp. https://doi.org/10.1093/oso/9780198767220.001.0001

Timana-Mendoza C, Reyes-Calderón A, Venail P, Britzke R, Santa-Maria MC, Araújo-Flores JM, Silman M, Fernandez LE (2025) Hydrological Connectivity Enhances Fish Biodiversity in Amazonian Mining Ponds: Insights from eDNA and Traditional Sampling. Molecular Ecology 17784(11): e17784. https://doi.org/10.1111/mec.17784

Zuanon J, Medonça FP, Espírito Santo H, Dias MS, Galuch AV, Akama A (2015) Guia de peixes da Reserva Adolpho Ducke. Editora INPA, 157 pp. https://repositorio.inpa.gov.br/handle/1/4705

Supplementary material 1

Reads abundance and number of ASV (haplotype) for all taxa (leaves of Fig. 2) identified across the 12 studied sites

Authors: Guillain Estivals, Ronald Delgado-Barboza, Morgan Ruiz-Tafur, Junior Chuctaya, Pierre Caminade, Carmen Garcia-Davila, Nicolas Hubert

Data type: xlsx

Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: https://doi.org/10.3897/zookeys.1270.170412.suppl1

Supplementary material 2

List of the 131 unassigned ASV obtained with the Lowest Common Ancestor (LCA) in Megan, with for each of them the three top scores and the closest species (identity <98%)

Authors: Guillain Estivals, Ronald Delgado-Barboza, Morgan Ruiz-Tafur, Junior Chuctaya, Pierre Caminade, Carmen Garcia-Davila, Nicolas Hubert

Data type: xlsm

Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: https://doi.org/10.3897/zookeys.1270.170412.suppl2

Supplementary material 3

List of fasta sequences of the 226 ASVs identified

Authors: Guillain Estivals, Ronald Delgado-Barboza, Morgan Ruiz-Tafur, Junior Chuctaya, Pierre Caminade, Carmen Garcia-Davila, Nicolas Hubert

Data type: fasta

Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: https://doi.org/10.3897/zookeys.1270.170412.suppl3